

Singapore Management University Institutional Knowledge at Singapore Management University

Research Collection School Of Economics

School of Economics

7-2014

Bayesian Analysis of Bubbles in Asset Prices

Andras FULOP
ESSEC Business School

Jun YU
Singapore Management University, yujun@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/soe_research



Part of the [Econometrics Commons](#), and the [Finance Commons](#)

Citation

FULOP, Andras and YU, Jun. Bayesian Analysis of Bubbles in Asset Prices. (2014). 1-35. Research Collection School Of Economics.
Available at: https://ink.library.smu.edu.sg/soe_research/1589

This Working Paper is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Bayesian Analysis of Bubbles in Asset Prices

Andras Fulop and Jun Yu

July 2014

Paper No. 04-2014

Bayesian Analysis of Bubbles in Asset Prices*

Andras Fulop and Jun Yu

July 17, 2014

Abstract

We develop a new asset price model where the dynamic structure of the asset price, after the fundamental value is removed, is subject to two different regimes. One regime reflects the normal period where the asset price divided by the dividend is assumed to follow a mean-reverting process around a stochastic long run mean. This latter is allowed to account for possible smooth structural change. The second regime reflects the bubble period with explosive behavior. Stochastic switches between two regimes and non-constant probabilities of exit from the bubble regime are both allowed. A Bayesian learning approach is employed to jointly estimate the latent states and the model parameters in real time. An important feature of our Bayesian method is that we are able to deal with parameter uncertainty, and at the same time, to learn about the states and the parameters sequentially, allowing for real time model analysis. This feature is particularly useful for market surveillance. Analysis using simulated data reveals that our method has better power for detecting bubbles compared to existing alternative procedures. Empirical analysis using price/dividend ratios of S&P500 highlights the advantages of our method.

Keywords: Parameter Learning, Markov Switching, MCMC

JEL Classification: C11, C13, C32, G12

*Fulop is with Finance Department, ESSEC Business School, Paris-Singapore, Cergy-Pontoise Cedex, France 95021. E-Mail: fulop@essec.fr; Yu is with School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903. E-Mail: yujun@smu.edu.sg. Yu would like to acknowledge the financial support from Singapore Ministry of Education Academic Research Fund Tier 2 under the grant number MOE2011-T2-2-096.

1 Introduction

The recent global financial crisis and the ongoing European debt crisis have prompted economists and regulators to work arduously to find ways to avoid the next crisis. From a historical perspective, Ahamed (2009) argues that financial crises are often preceded by an asset market bubble.¹ Well-known bubble episodes include the Dutch tulip mania, the British South Sea bubble, the Railway mania, Roaring Twenties stock-market bubble, the Dot-com bubble, and the US housing bubbles related to the subprime mortgage. All these bubbles were followed by financial crises. Bubbles are generally considered harmful to economics and the welfare of society and they lead to the mis-allocation of resources. In a recent article, Caballero, Farhi and Gourinchas (2008) argues the burst of an asset price bubble can lead to recession in real economy. Consequently, approaches have been tried to detect the presence and the burst of financial bubbles and to estimate the bubble origination and collapsing dates.

Broadly speaking there are three alternative models in the bubble literature. The first approach employs the regime switching models. The second approach is based on various structural break models. The last approach is based on noncausal processes.

The regime switching models have a long history in economics, dating back to Goldfeld and Quandt (1973) and Hamilton (1989). Evans (1991)'s model may be regarded as a regime switching model with two regimes. One regime corresponds to bubble expansion whereas the other regime corresponds to bubble collapse. The collapse is sudden and takes place within a single period. After the bubble collapses, a new bubble starts to emerge. The collapse of the bubble is determined by an independent Bernoulli trial. Another regime switching model was proposed by Funke, Hall and Sola (1994) and Hall, Psaradakis and Sola (1999). In this model, two regimes have been used: one regime has a unit root and the market is efficient, whereas the other regime has a rational bubble and hence has an explosive root. More recently, Shi (2013) extends the model

¹As Federal Reserve Board former vice chairman – Donald Kohn – argued, Federal Reserve policymakers should deepen their understanding about how to combat speculative bubbles to reduce the chances of another financial crisis.

in Hall, Psaradakis and Sola (1999) to allow for heteroskedasticity. Shi and Song (2014) proposed to use an infinite hidden Markov model, which allows for infinite number of regimes, to detect, date stamp, and estimate speculative bubbles.

The structural break models have been extensively used to distinguish stationary models and unit root models; see for example, Kim (2000) and Buseti and Taylor (2004). Recently, Homm and Breitung (2012) extends some of the methods to distinguish explosive models and unit root models. The models considered in Homm and Breitung (2012) have one change point. On one side of the change point, there is a unit root in the dynamic structure. On the other side of the change point, the model has an explosive root. The change point is not stochastic in these models. This deterministic nature of modelling structural change point seems more restrictive than the regime switching models.

More recently, Gouriéroux and Zakoian (2013) proposes to use a noncausal Cauchy autoregressive process to model explosiveness and showed that the model can explain multiple bubbles phenomenon although the model is a strictly stationary process.

Regarding statistical inference of the presence of bubbles and the date-stamping of bubble origination and termination, several methods have been proposed. The first method is based on the full sample maximum likelihood (ML) method. This includes Funke, Hall and Sola (1994) and Hall, Psaradakis and Sola (1999) in the context of regime switching models. When the model is correctly specified, the ML estimator (MLE) is efficient. Probabilistic inference about the unobserved regimes can be based on the Hamilton filter by calculating either the filter probability or the smoothed probability. These probabilities naturally depend on the unknown parameters. To estimate the probabilities, the unknown parameters are replaced by the MLE obtained from the full sample. Consequently, the inferential approach does not allow for real time analysis as there is no sequential learning about the parameters. This feature of lack of real time analysis is shared by some MCMC algorithms in the literature such as the one in Shi and Song (2014).

The second method is based on recursive techniques. For example, Phillips, Wu

and Yu (2011) suggests implementing the right-tailed ADF test repeatedly on a forward expanding sample sequence. To effectively deal with episodes with multiple bubbles, Phillips, Shi and Yu (2013a, 2013b, PSY hereafter) varies both the initial point and the ending point of the sample in each recursive regression. Hogg and Breitung (2012) modifies various recursive methods for the purpose of bubble detection and date-stamping of bubble origination and termination. Apart from the ease in implementation, a nice feature of the recursive method is that it provides real time estimate of the bubble state. In practice, however, it is possible that the chosen minimum window size is larger than the actual bubble duration. Moreover, for the test statistic to rise above the critical value, a long enough period and a strong enough signal from the explosive regime are needed. Not surprisingly, in finite sample, the method may overestimate the bubble origination and collapsing dates.

In this paper, we make several contributions to the empirical asset pricing literature. First, we propose a two-state regime switching model of bubbles that generalizes the existing literature. One state reflects the normal period corresponding to a common stochastic trend between asset prices and dividends. Given that the underlying series that we model is a price divided by a fundamental (e.g. a price-dividend or P/D ratio) we assume a mean-reverting dynamics in this state around a potentially time varying long-run mean. One can think of this mean-reverting behavior as a result of low-frequency cyclical movements in the discount rates suggested in the recent finance literature (Cochrane, 2011). In addition, to allow for smooth permanent structural changes in asset markets, we allow the long-run mean of the series itself to follow a random walk process. Evidently when the variance of this latter is set to zero, smooth structural change is excluded. For the duration of this normal regime we assume a standard exponential distribution that implies Markovianity of the regime changes. The second state reflects the bubble period where the AR coefficient is larger than 1. Here we depart from the extant regime-switching literature and allow for non-constant hazard rates of exit from the bubble regime corresponding to Weibull-distributed durations. Throughout we assume normally distributed innovations with heteroskedasticity across

regimes. Our second contribution is to implement a Bayesian learning approach for making sequential joint statistical inference over latent states, model parameters, and model comparison. There are several appealing features of this new inferential method for bubble detection. Firstly, based on the regime switching method, we can avoid the need of specifying the minimum duration of each regime, including the bubble regime. In PSY, the minimum duration is set to the minimum window size of regression. In the case when the minimum window size has to be specified, if the minimum window size is larger than the minimum duration of a regime, the identification of a regime will be biased. Consequently, it is reasonable to believe that our model can identify the change points more effectively and more quickly when there are quick regime shifts. Secondly, the change points are endogenously determined. Thirdly, we are able to deal with parameter uncertainty as well as learning about the states and the parameters sequentially, allowing for real time model analysis. This feature is particularly useful for market surveillance, as argued in PSY (2013a). Fourthly, our approach enables the exact finite sample inference about the parameter as well as latent states and hence avoids the derivation of asymptotic distribution. As shown in PSY, the asymptotic properties can be very difficult to obtain in general and this is especially true for the estimator of the change point.

To check the reliability of the proposed method for the model, we conduct a Monte Carlo study. The Monte Carlo results show that the method is reliable both for the estimation of parameters and more importantly for detecting bubbles. Comparing its performance to PSY we find that our Bayesian learning algorithm reacts faster and has better power to detect bubbles. To investigate the robustness of the method with respect to outliers, we also simulate data from a model with fat tailed innovations but estimate the (misspecified) model without fat tails using the proposed method and PSY and find that both methods are robust. We also apply our method to real data, i.e., monthly S&P 500 P/D ratio between 1871 and 2012, as PSY did.

Our empirical estimates of the timing of bubbles are broadly similar to the empirical results obtained by PSY (2013a) based on a recursive frequentist method. However,

our procedure flags more bubble episodes than PSY and is more capable to differentiate between bubble periods and normal periods. Furthermore, in line with our simulation evidence, we find that a decision maker who is averse to erroneously declaring a regime change will identify substantially fewer and longer bubble periods.

The organization of the paper is as follows. Section 2 introduces the model and proposes a new estimation method. Section 3 presents simulation results. Section 4 reports the empirical results and Section 5 concludes.

2 Econometric Model and Estimation Method

2.1 Model and inferential task

We assume the presence of two regimes determining the autoregressive behavior of the series, with $s_t = 1$ the normal regime, $s_t = 2$ the bubble regime. The distribution of the duration of the normal regime is exponential with parameter λ_1 , i.e., if the duration of the normal spell is denoted by τ_n , then we have $Pr(\tau_n > t) = \exp(-t/\lambda_1)$. As our focus is on the bubble regimes, we assume that the bubble duration, τ_b , follows the more flexible Weibull distribution with parameters λ_2, k_2 , giving rise to the survival probabilities

$$Pr(\tau_b > t) = \exp\left(-(t/\lambda)^{k_2}\right).$$

The expected value of the bubble spell is $\mu_2 = \lambda_2\Gamma(1+1/k_2)$ and we reparameterize the model and define our prior over μ_2 instead of λ_2 . The shape parameter k_2 determines whether the hazard rate of exit is constant, increasing or decreasing.

The process in each regime follows

$$s_t = 0 \quad : \quad x_t = \alpha_t(1 - \beta_1) + \beta_1 x_{t-1} + \sigma_t \varepsilon_t, \beta_1 < 1, \quad (1)$$

$$\alpha_t = \alpha_{t-1} + \delta \eta_t, \quad (2)$$

$$s_t = 1 \quad : \quad x_t = \beta_2 x_{t-1} + \sigma_t \varepsilon_t, \beta_2 > 1. \quad (3)$$

In the normal state (1), x_t follows a mean-reverting process around the stochastic mean α_t , where the speed of mean-reversion is β_1 . To allow for gradual parameter change in the long-run mean, Equation (2) posits that α_t follows a random walk whose variability is determined by δ . Obviously, when $\delta = 0$, we are back to a constant mean reversion model. In the state with an explosive root (3), we claim that there is a bubble in the asset price. This is because we will understand x_t as an asset price with the fundamental value removed. As a result, the presence of an explosive root implies the presence of bubble according to the present value model; see, for example, Diba and Grossman (1988). Following the suggestion of Phillips, Wu and Yu (2011), we do not use an intercept in the explosive state for otherwise the intercept would dominate the autoregressive term asymptotically which is not empirically realistic.²

To address the concern of Shi (2013) about the sensitivity to bubble identification to the presence of heteroskedasticity, we allow σ_t to follow an independent 2-regime Markov switching process, with diagonal probabilities z_{ii} . In the first (low) regime the value of volatility is $\sigma_t = \sigma_l$ while in the second (high) regime, $\sigma_t = \sigma_m \sigma_l$ where $\sigma_m > 1$. The transition matrix for volatility is

$$P_\sigma = \begin{bmatrix} z_{11} & 1 - z_{22} \\ 1 - z_{11} & z_{22} \end{bmatrix}.$$

The fixed parameter vector describing the dynamics of the system has 10 unknown parameters $\theta = (\lambda_1, k_2, \mu_2, z_{11}, z_{22}, \sigma_l, \sigma_m, \delta, \beta_1, \beta_2)'$.

To monitor bubbles in real time, the user of the model needs to evaluate the probability of being in a bubble (or normal) regime at time t , given information available by time t . Even if he knows the fixed parameters θ , inference over the regimes, i.e. obtaining

$$p(s_t = k \mid \theta, x_{1:t}) = E(1_{\{s_t=k\}} \mid \theta, x_{1:t}),$$

is not easy as the filter is not analytically available for the model. However, in what

²In a more recent attempt, PSY (2014) show the impact of the intercept term on the asymptotics.

follows we describe a very efficient sequential Monte Carlo technique (a particle filter) to numerically approximate the filtering distributions.

2.2 Discrete particle filter

The theoretical quantity that the filtering algorithm targets is the sequence of filtering distributions of the state-space system $f(s_t, h_t, \sigma_t, \alpha_t \mid x_{1:t}, \theta)$ where h_t is the time elapsed since the last regime change. Throughout this section we assume a known parameter vector θ and, to simplify notation, we suppress dependence on it. The crucial thing to realize is that conditional on the path of the discrete latent states $s_{1:t}, h_{1:t}, \sigma_{1:t}$ the system is a linear Gaussian state space model, and hence the continuous state variable α_t can be marginalized out analytically using Kalman filtering recursions. Let us denote the two filtering moments of α_t (conditional on discrete latent variables) by μ_t, V_t , and hence the joint filtering density to track becomes $f(s_t, h_t, \sigma_t, \mu_t, V_t \mid x_{1:t})$. Given that the state space that we need to filter numerically is discrete, we can employ the discrete particle filter (DPF) of Fearnhead (1998) where all successor states are generated avoiding the use of a proposal distribution. Let us assume that at $t - 1$ we have N equal-weighted particles $(s_{t-1}^i, h_{t-1}^i, \sigma_{t-1}^i, \mu_{t-1}^i, V_{t-1}^i), i = 1, \dots, N$ with weights $s_{t-1}^i = \frac{1}{N}$ representing the filtering distribution. We have the following recursion to arrive to the filtering distribution at the next time instant t .

Branching out: To move the hidden state particles forward, one needs to attach s_t, σ_t to the existing particles to characterize $f(s_t, \sigma_t, s_{t-1}, h_{t-1}, \sigma_{t-1}, \mu_{t-1}, V_{t-1} \mid x_{1:t-1})$. Instead of some random proposal over the new states the DPF proposes to create all possible successor states from each existing particles. In our case we have $K = 4$ possible successor particles for each ancestor corresponding to all possible configurations of s_t, σ_t . This results in $4 \times N$ particles of the form $(s_t = k, \sigma_t = l, s_{t-1}^i, h_{t-1}^i, \sigma_{t-1}^i, \mu_{t-1}^i, V_{t-1}^i), k = 0, 1; l = 1, 2, i = 1, \dots, N$ with attached weights

$$s_{t|t-1}^{ikl} = f(s_t = k, \sigma_t = l \mid s_{t-1}^i, h_{t-1}^i, \sigma_{t-1}^i) s_{t-1}^i.$$

Attaching new information and computing the likelihood: Next, the particles are reweighted to include the effect of the new observation x_t . The theoretical relationship between the predictive distribution and the filtering one is

$$\begin{aligned} & f(s_t, \sigma_t, s_{t-1}, h_{t-1}, \sigma_{t-1}, \mu_{t-1}, V_{t-1} \mid x_{1:t}) \\ \propto & f(x_t \mid s_t, \sigma_t, \mu_{t-1}, V_{t-1}) f(s_t, \sigma_t, s_{t-1}, h_{t-1}, \sigma_{t-1}, \mu_{t-1}, V_{t-1} \mid x_{1:t-1}). \end{aligned}$$

This can be implemented in the algorithm by reweighting to arrive to the filtering weights $\tilde{s}_t^{ikl} = f(x_t \mid s_t = k, \sigma_t = l, \mu_{t-1}^i, V_{t-1}^i) s_{t|t-1}^{ikl}$. The estimate of the marginal likelihood of x_t can be computed as

$$\hat{p}(x_t \mid x_{1:t-1}) = \frac{1}{4N} \sum_{k=1}^2 \sum_{l=1}^2 \sum_{i=1}^N \tilde{s}_t^{ikl}.$$

Resampling: Clearly, repeating the previous steps through multiple observations would lead to an exponential growth in the number of discrete states to be maintained. Hence it is crucial to include a resampling step where N particles are sampled out of the $4 \times N$ existing one with probability proportional to the normalized weights $s_t^{ikl} = \frac{\tilde{s}_t^{ikl}}{\sum_{k=1}^2 \sum_{l=1}^2 \sum_{i=1}^N \tilde{s}_t^{ikl}}$. This results in an N -sample, $(s_t^i, \sigma_t^i, s_{t-1}^i, h_{t-1}^i, \sigma_{t-1}^i, \mu_{t-1}^i, V_{t-1}^i), i = 1, \dots, N$ with equal weights $s_t^i = \frac{1}{N}$. The last step is to update the hidden variables (h_t^i, μ_t^i, V_t^i) which are simply deterministic functions of their past values, the new state variables s_t^i, σ_t^i and of the observation x_t .

The empirical distribution of the particle cloud converges to the true filtering density under weak conditions and it can be used to approximate any filtering quantity of interest. For example the filtered bubble probability can be approximated as:

$$p(s_t = 1 \mid \theta, x_{1:t}) \approx \hat{p}(s_t = 1 \mid \theta, x_{1:t}) = \frac{1}{N} \sum_{i=1}^N 1_{\{s_t^i=1\}}.$$

2.3 Parameter learning algorithm

In practice, we do not know θ . A common practice in the regime switching literature is to replace θ by the ML estimates or the Bayesian estimates obtained from the full sample, ignoring the parameter uncertainty. Since the estimates of θ are constructed from the full sample of the data, such an analysis is not in real time. To carry out a real time analysis, the model parameters also need to be sequentially updated as new data arrive. Furthermore, ignoring parameter uncertainty can lead to an overestimation of our ability to regime detection in real time, especially for more complex models.

To tackle these issues we turn to sequential Bayesian techniques that allow us to sample from the posterior probability of the fixed parameters $p(\theta \mid x_{1:t})$. As an example, assume that we have a weighted sample $(\pi_t^m, \theta_t^m, m = 1, \dots, M)$ with normalized weights $\sum_{m=1}^M \pi_t^m = 1$ whose empirical distribution approximates $p(\theta \mid x_{1:t})$. Further, assume that for each θ_t^m , we have N state particles $s^{m,i}, i = 1, \dots, N$ approximating $f(s_t \mid x_{1:t}, \theta_t^m)$ obtained by running a DPF at θ_t^m . Then the posterior probabilities that take parameter uncertainty into account can be computed as

$$E(1_{\{s_t=k\}} \mid x_{1:t}) = E(E(1_{\{s_t=k\}}, \theta, x_{1:t}) \mid x_{1:t}) \approx \sum_{m=1}^M \pi_t^m \frac{1}{N} \sum_{i=1}^N 1_{\{s_t^{m,i}=k\}}.$$

The use of Bayesian methods enables us to conduct the exact finite sample inference of the parameters as well as latent states and hence avoids the derivation of asymptotic distribution. The derivation of asymptotic properties of classical estimators can be very difficult to obtain for this class of models due to the presence of explosiveness. This difficulty is especially true for the estimator of the change point; see, for example, PSY (2013b).

Sequential analysis of state-space models under parameter uncertainty is of interest in many settings. Since one of our primary interests here is real time analysis of regime detection, parameter learning is needed. To sequentially learn over the parameters, we turn to the marginalized resample-move approach of Fulop and Li (2013) and Chopin et al. (2013). For completeness, we provide a brief overview of the method in this

subsection. We need a method to sequentially sample from the sequence of posteriors

$$\gamma_t(\theta) = p(\theta \mid x_{1:t}) \propto p(x_{1:t} \mid \theta)p(\theta) = \prod_{l=1}^t p(x_l \mid x_{1:l-1}, \theta)p(\theta), \quad (4)$$

where $p(\theta)$ is the prior over the fixed parameters and for notational convenience we suppress the dependence on the initial hidden state. While the individual conditional likelihoods $p(x_l \mid x_{1:l-1}, \theta)$ cannot be obtained in closed form, Fulop and Li (2013) and Chopin et al (2013) employ instead the approximate likelihoods obtained from particle filters. In particular, instead of the targets in (4) they propose to work with the extended target

$$\tilde{\gamma}_t(\theta, u_{1:t}) = \prod_{l=1}^t \hat{p}(x_l, u_l \mid x_{1:l-1}, u_{1:l-1}, \theta) \phi(u_l \mid x_{1:l}, u_{1:l-1}) p(\theta), \quad (5)$$

where the likelihood estimates $\hat{p}(x_l, u_l \mid x_{1:l-1}, u_{1:l-1}, \theta)$ (denoted as $\hat{p}(x_t \mid x_{1:t-1})$ in the previous section) are obtained by running a particle filter with N particles for any given fixed parameter θ , u_l contains all the random variables created at time l by the particle filter and $\phi(u_l \mid x_{1:l}, u_{1:l-1})$ is the density of these random variables.

Initialization: Sample particles from the prior $\theta_0^m \sim p(\theta)$, and attach equal weights to each particle $\pi_0^m = \frac{1}{M}$. The resulting cloud, (π_0^m, θ_0^m) is trivially distributed according to $p(\theta)$. For each θ_0^m initialize a particle filter with N state particles and denote the random variables created by u_0^m .

Recursion and reweighting: Assume that a weighted sample, $(\pi_{t-1}^m, \theta_{t-1}^m)$, has been obtained that represents $p(\theta \mid x_{1:t-1})$. Furthermore, for each θ_{t-1}^m we maintain a particle filter with N state particles with attached random variables $u_{1:l-1}^m$ and the likelihood estimates up to $t-1$, $\hat{p}(x_{1:t-1} \mid \theta_{t-1}^m)$. Now the task is to include the new observation into the information set and obtain a representation of the next posterior, $p(\theta \mid x_{1:t})$. The sequential resample-move algorithm uses importance sampling for this task and the intuition that the posterior at $t-1$ is typically quite close to the posterior at t . Hence, the sample from the former will provide a good proposal distribution for the latter.

Then the location of the particles is inherited from the previous cloud: $\theta_t^m = \theta_{t-1}^m$. However, the importance weights will change to account for the difference between the target and the proposal leading to new un-normalized weights

$$w_t^m = \pi_{t-1}^m \times \frac{\tilde{\gamma}_t(\theta_t^m, u_{1:t}^m)}{\phi(u_t^m \mid x_{1:t}, u_{1:t-1}, \theta_t^m) \tilde{\gamma}_{t-1}(\theta_{t-1}^m, u_{1:t-1}^m)} = \pi_{t-1}^m \times \hat{p}(x_t, u_t^m \mid x_{1:t-1}, u_{1:t-1}^m, \theta_t^m), \quad (6)$$

The incremental weights $\hat{p}(x_t, u_t^m \mid x_{1:t-1}, u_{1:t-1}^m, \theta_t^m)$ are obtained by running the particle filter for each θ_t^m on the new observation x_t and recording the resulting likelihood estimate. The random numbers used in the particle filters need to be independent across θ_t^m and through time. The new normalized weights are $\pi_t^m = \frac{w_t^m}{\sum_{i=1}^M w_t^i}$ and the particle cloud, (π_t^m, θ_t^m) , will represent the target $p(\theta \mid x_{1:t})$.

Sample Degeneracy: If we keep repeating the reweighting steps, at some point the sample would degenerate. We measure sample diversity by the Efficient Sample Size: $ESS = \frac{1}{\sum_{m=1}^M (\pi_t^m)^2}$. Whenever the ESS drops below some prespecified number B , we will introduce additional algorithmic steps to improve the support of the distribution.

Resampling Step: First, to focus computational efforts on the more likely part of the sample space, we resample the particles proportional to weights π_t^m to arrive to an equal-weighted sample and set weights to $\pi_t^m = 1/N$. Any resampling method can be used where choice probabilities are proportional to the weights. Also, notice that for each resampled θ^m , the attached dynamic states and likelihood estimates need to be resampled too. The resulting cloud is still distributed according to $\tilde{\gamma}_t(\theta, u_{1:t})$.

Move Step: The resampling step in itself does not enrich the support of the particle cloud. We need to boost the particle cloud in such a way that does not distort its distribution in a probabilistic sense as M goes to infinity. This can be achieved by moving each particle through a kernel $K_t(\cdot \mid \theta, u_{1:t})$ that admits $\gamma_t(\theta, u_{1:t})$ as an equilibrium distribution. $K_t(\cdot \mid \theta)$ will be a particle marginal M-H kernel from Andrieu et al (2010):

- Propose from a proposal density: $(\theta^*, u_{1:t}^*) \sim \phi(u_{1:t}^* \mid \theta^*) h_t(\theta^* \mid \theta^i)$.

- Acceptance probability:

$$\begin{aligned}\alpha &= \min \left(1, \frac{\tilde{\gamma}_t(\theta^*, u_{1:t}^*) \phi(u_{1:t}^i | \theta^i) h_t(\theta^i | \theta^*)}{\tilde{\gamma}_t(\theta^i, u_{1:t}^i) \phi(u_{1:t}^* | \theta^*) h_t(\theta^* | \theta^i)} \right) \\ &= \min \left(1, \frac{\widehat{p}(x_{1:t}, u_{1:t}^* | \theta^*) h_t(\theta^i | \theta^*)}{\widehat{p}(x_{1:t}, u_{1:t}^i | \theta^i) h_t(\theta^* | \theta^i)} \right).\end{aligned}$$

- With probability α , set $(\theta^i, u_{1:t}) = (\theta^*, u_{1:t}^*)$, otherwise keep original value.

Here, $h_t(\cdot | \theta^{(n)})$ can be fitted on the cloud of particles to better approximate the target. Fulop and Li (2013) and Chopin et al (2013) show that this algorithm actually delivers exact inference over the sequence of joint filtering distribution of the parameters and the dynamic states.

Notice that the move step needs to browse through the full data-set $x_{1:t}$ as the likelihood is evaluated at each new proposal θ^* . In contrast, the reweighting step only needs the last individual likelihood $\widehat{p}(x_t | x_{1:t-1}, \theta)$. Hence the time needed for a move step keeps increasing with the sample size, while the reweighting step has a constant speed. Fortunately, as the sample becomes large the posteriors stabilize and one needs to resort to move steps less and less often. To keep the estimation error in the likelihood under control, under some mixing assumption on the state-space model the number of state particles need to increase linearly with overall sample size. Chopin et al (2013) show that overall the computational cost of the algorithm is of the order t^2 .

2.4 Loss functions for bubble-stamping

The discrete decision the policy maker is faced at time t is whether he stamps the given period as a bubble or not. Let us denote this decision as $a_t = 1$ if the period is stamped as a normal period, $a_t = 0$ if it is stamped as a bubble period. Then the loss function related to the decision is defined as

$$L_t(s_t, a_t) = l_t^b a_t 1_{s_t=0} + l_t^n (1 - a_t) 1_{s_t=1}, \quad (7)$$

where l_t^b is the loss from classifying a normal period as a bubble period while l_t^n is the loss from classifying a bubble period as a normal period. If the policy maker has a preference of avoiding “crying wolf”, we can let these loss functions state-dependent, i.e., if the previous period was stamped as a normal period ($a_{t-1} = 0$), we would have $l_t^b > l_t^n$ and the reverse if the previous period was stamped as a bubble period. The expected loss of the policy maker based on real time information corresponding to the two decisions are

$$E_t(L_t(s_t, a_t = 1)) = l_t^b Pr(1_{\{s_t=0\}} \mid x_{1:t}), \quad (8)$$

$$E_t(L_t(s_t, a_t = 0)) = l_t^n Pr(1_{\{s_t=1\}} \mid x_{1:t}). \quad (9)$$

Hence, flagging the period as a bubble period is optimal if

$$\frac{l_t^b}{l_t^n} < \frac{Pr(1_{\{s_t=1\}} \mid x_{1:t})}{Pr(1_{\{s_t=0\}} \mid x_{1:t})}. \quad (10)$$

Notice that for decision-making it is only the ratio of the loss functions denoted by $\zeta = \frac{l_t^b}{l_t^n}$ that matters. If we assume that this parameter is time invariant conditional on being in a regime stamped as a normal regime, the rule is to stamp the new regime as a bubble if $\zeta < \frac{Pr(1_{\{s_t=1\}} \mid x_{1:t})}{Pr(1_{\{s_t=0\}} \mid x_{1:t})}$. By analogy, we stamp the end of a bubble spell whenever $\zeta < \frac{Pr(1_{\{s_t=0\}} \mid x_{1:t})}{Pr(1_{\{s_t=1\}} \mid x_{1:t})}$.

3 Monte Carlo Study

3.1 Priors and parameter restrictions

In the simulation study and the empirical study, we use the following priors for model parameters:

- λ_1 : This is the parameter determining the expected length of a normal regime. In particular we use a normal distribution truncated from below at 120 months to reflect our a priori beliefs that the normal regime should be reasonably long

lasting. The mean and standard deviation parameters of the normal are 180 and 60 respectively.

- k_2 : Determines the shape of the bubble regime distribution. Here we assume that k_2 has an increasing probability of exit from the bubble state in duration, so assume a normal distribution truncated from below with mean and standard deviation parameters of the normal of 1.
- μ_2 : Determines the expected length of a bubble spell. Here again we use an informative prior to focus on reasonably long-lasting bubble spells. In particular, we use a normal distribution truncated from below at 24. The mean and standard deviation parameters of the normal are 36 and 12 respectively.
- z_{11}, z_{22} : For both we use a uniform prior on $[0, 1]$.
- σ_l : We use a normal prior truncated below at 0, with $\mu = std(\Delta x_t), \sigma = std(\Delta x_t)$.
- σ_m : We use a normal prior truncated below at 1, with $\mu = 1, \sigma = 1$.
- δ : We use a normal prior truncated below at 0, with $\mu = 0, \sigma = 0.25$.
- β_1 : Measures the mean reversion during the normal regime. The recent literature in finance points towards time-varying discount rates as the main source behind the cyclical variation in the P/D ratio (see for instance Cochrane (2011) for a recent overview), usually thought of as a medium-to-low frequency phenomenon. Hence we bound from below the half-life of the mean reversion at 2 years, corresponding to $\beta_1 = 0.9715$ with monthly data. In addition to this we only assume non-explosiveness of the process in the normal regime leading to the uniform prior $\beta_1 \sim U[0.9715, 1]$.
- β_2 : We simply use a $\beta_2 \sim U[1, 1.02]$ prior where the upper boundary is chosen to make sure that we cover all empirically relevant parameters of β_2 .

In the SMC procedure we use $M = 2,048$ parameter particles with $N = 128$ state particles in each particle filter. The resample-move step is triggered when the efficient

sample size drops below $N/2$. In the move-steps we use an independent mixture normal proposal. The routine has been coded in MATLAB with the particle filtering operation programmed efficiently in CUDA and run on a Kepler K20 GPU.

3.2 Monte Carlo results

To investigate the reliability of the learning routine we design two Monte Carlo experiments. In the first experiment, we simulate 100 data sets from the proposed regime switching model, each with $T = 1,698$ observations. The parameter setting, including the sample size, is similar to what has been found in the empirical study from S&P500. The first column of Table 1 reports the parameter values used to generate the data. The second column reports the average of the full-sample posterior means and the third column the average length of the central posterior 90% credible interval. For comparability the last column shows the prior central 90% credible interval. The results show that the informativeness of the data varies starkly across the various parameters. First, one can see that the length of the posterior credible interval for λ_1, k_2, μ_2 is only a bit smaller compared to the prior analogue, mirroring limited learning about these parameters that determine the regime transition probabilities. This is a natural consequence of the small average number of regime changes in the simulated data. Not surprisingly, the average posterior means are markedly biased towards the prior means. A similar phenomenon can be observed for δ that controls the variability of the long-run mean of the time-series. The reason here is that changes in α_t happen at a low frequency resulting in a small effective sample. The remaining parameters, $z_{11}, z_{22}, \sigma_l, \sigma_m, \beta_1, \beta_2$, describe time-varying volatility and the within-regime conditional means and are much better identified from the data. The posterior means are close to the real generating values and the posterior credible intervals are a fraction of the prior ones.

Now we turn to the ability of our model to detect explosive periods (bubbles) in the underlying time series and compare it with the real time detection algorithm in PSY. In our Bayesian algorithm we investigate several values of the date-stamping parameter ζ . As a baseline, we look at the case of symmetric loss, i.e. $\zeta = 1$. Then, we also investigate

cases where the decision maker is averse to changing the stamp too quickly. Specifically, we look at $\zeta = 2$ and $\zeta = 3$. For PSY, we follow Yiu, Yu and Jin (2013) and declare a switch to a bubble stamp whenever the backward sup ADF (BSADF) test statistic surpasses $Q_{BSADF}(0.95) + \log(t)/100$ where $Q_{BSADF}(0.95)$ is the 95% critical value of the test statistic obtained by Monte Carlo simulation. Further, we stamp the end of the bubble when the statistic drops below $Q_{BSADF}(0.95)$. Note that both algorithms are implementable using only data available in real time. In the upper panel of Figure 1 we plot with solid lines the detection rates of our Bayesian learning algorithm as a function of bubble duration, where the flat line corresponds to $\zeta = 1$, the line with circles to $\zeta = 2$ and the line with plus signs to $\zeta = 3$. The dashed line reports the detection rates from using the BSADF statistic as in PSY. All results are averaged across 100 data sets using the true DGP to simulate the data. One can see that our Bayesian methodology seems to react faster with better detection rates when only a few periods has elapsed since the start of the bubble. The power of the two algorithms seems to get closer as the length of bubble becomes large. Importantly, the detection capability of the test does not seem to deteriorate much with an increase in the bubble stamping parameter ζ . Let us note that the size of the Bayesian test (periods flagged as bubble that are in fact not bubbles) is around 0.018 while for PSY it is 0.057. There are at least two usual caveats with the use of relatively richly parameterized nonlinear models like our regime-switching one in time-series: First, one is concerned about parameter uncertainty, second about the robustness of the results to the exact model specification. As mentioned earlier, our Bayesian learning approach deals with the first of these as it fully takes parameter uncertainty into account. However, one may still wonder about the extent to which these detection results are conditional on the model being exact. In particular, given the well-known fact that asset returns are non-normal, in the lower panel of Figure 1 we investigate the detection capability of our method and that of PSY in a setting where the data innovations follow a fat-tailed student-t distribution with 4 degrees of freedom instead of the normal postulated by our regime switching model. The results closely mirror the ones in the upper panel showing the robustness of both

methods to the presence of fat-tailed innovations.

To gain further insight into the behavior of identified bubbles by the alternative bubble indicators. Table 2 reports some summary statistics of identified bubbles, including the number of identified bubble episodes, the proportion of the identified bubble periods, and the average bubble duration, both for PSY and the regime switching model with different value of ζ . The first and third columns show that both PSY and the regime switching model with $\zeta = 1$ (no aversion to regime change) indicate more frequent but shorter bubbles compared to $\zeta > 1$. Hence, allowing for a "regime-change-averse" loss function for the decision maker provides a principled way to flag fewer and longer bubbles in the regime-switching framework. Furthermore, the second column indicates that the proportion of periods labeled as bubble is much less variable across the different methods. Hence, these results are not simply the result of using a more aggressive test procedure.

4 Empirical Study

In the empirical study, we confront the proposed model with a well-known time series. We fit the model to the monthly real S&P500 series and dividend series as in PSY (2013a). The data series is from January 1871 to June 2012, resulting in 1,698 observations and downloadable from Professor Shiller's website. As in PSY, we investigate the series of the price-dividend ratio, plotted in Figure 2. The choice of the data over a long time span reflects our objective to capture as many stock market phases as possible. In the meantime, the use of the same data as in PSY allows us to compare our estimates with those of PSY.

Table 3 presents the full sample posterior estimates obtained from our learning routine and Figure 3 shows the histograms of the priors (in red) alongside the full-sample posteriors (in blue). We can observe that while there is a large amount of uncertainty remaining about the parameters driving the regime changes, the data does tell us something about these parameters. First, the posteriors of both λ_1 and μ_2 put relatively

more weight onto lower values, suggesting somewhat more frequent regime changes compared to our priors. Second, the posterior over the shape parameter of the bubble duration k_2 seems to have a density separated away from 1 providing evidence against a simple exponential bubble duration distribution (case of $k_2 = 1$). This evidence points towards an increasing hazard function of exit from the bubble state, i.e. the probability of a crash tends to increase as bubbles mature. This is in contrast to a standard homogenous continuous time Markov-Switching model that gives rise to exponentially distributed durations. Further, the presence of stochastic volatility is clear in the data. The volatility in the high volatility regime is almost three time as large as that in the low volatility regime but is markedly less persistent. As in the Monte Carlo simulations, the data does not reveal too much about δ , the volatility of the long-run mean, but the posterior seems separated from zero. In contrast, the mean-reversion coefficient in the normal regime β_1 seems well identified with a mode that is close to but separate from unity and a posterior mean of 0.99. To translate this parameter to a more intuitive scale, we compute the posterior mean of the half-life of the process during the normal regime. If $\beta_1^i, i = 1, \dots, M$ are the posterior draws of the parameter, the posterior mean half-life is computed as $\widehat{HL}_{normal} = \frac{1}{M} \sum_{i=1}^M \frac{\ln 0.5}{\ln \beta_1^i}$. In our data set this results in an estimate of $\widehat{HL}_{normal} = 130$ i.e. 10.8 years. Last, the autoregressive coefficient during explosive regimes is tightly identified and symmetrically distributed around a posterior mean estimate of $\widehat{\beta}_2 = 1.015$.

Our main object of interest is not the parameter estimates *per se* but the ability to detect bubbles in real time using the filtered bubble probabilities. Table 4 reports some summary statistics on bubble-stamping both from PSY (2013a) (row 1) and from our algorithm with different values of ζ . The first thing to note is that in accordance with the simulation results, both PSY (2013a) and the regime switching model with $\zeta = 1$ seems to detect lots of short bubbles with the average bubble spell around 4 months in both cases. Further, increasing the date stamping parameter to $\zeta > 1$ seems to lead to an increase in the average length of the detected bubble spells while decreasing the number of bubble periods. Overall, allowing for aversion to change in the bubble

stamps leads to more intuitive results at least for this data set. Second, the results are not too different across $\zeta = 2$ and $\zeta = 3$. Hence, the results seem reasonably robust to the exact choice of the loss function. Last, the regime switching labels more periods as bubbles compared to PSY. Figure 4 shows the results from a real time bubble classifier with $\zeta = 2$ together with the filtered bubble regime probabilities. Here we label a given month a bubble (shaded in grey) if the bubble regime has the highest filtered probability. For comparison, we also implement the real time bubble indicator using the BSADF statistics from PSY in Figure 5. It is reassuring to observe that there are quite a few periods where the incidence of bubbles is preponderant according to both methods. In particular, around 1880, the years before 1920, before the great depression in 1929, the internet bubble before 2000 and last, the rebound after the recent 2008 financial crisis.

To better understand the behavior of the various bubble indicators, we take a microscopic view around well-known historical events. Here we focus on five events: The banking crisis in October 1907, the great market crash in September 1929, the Black Monday crash in October 1987, the DotCom mania peaking in March 2000 and the sub-prime crisis exploding in September 2008. Figure 6 and Figure 7 reports both the PSY BSADF statistic (row 1), the filtered bubble (row 2) probabilities our regime-switching model in the two years preceding and following these events. For reference row 3 depicts the original data series (P/D ratios) in the same periods. There are a few interesting patterns emerging from these graphs. First, both methods seem to indicate the presence of bubbles before the 1929 and 1987 crashes and during the DotCom bubble before 2000. A slight difference is that the regime-switching model seems to give more indication to the consecutive arrival of several shorter explosive periods, especially during the DotCom Mania. Second, the PSY method seems to have a difficulty in differentiating bubbles from collapses, a feature also noted in PSY (2013a, Footnote 28). For instance, the BSADF statistic takes large positive values during the market collapse before October 1907 or in the months right after the Lehman bankruptcy in 2008. In contrast in the regime switching model the bubble regime probabilities stay

low during these times. Third, the two methods interpret very differently when the market rallies after collapsing. For example, in months after October 1907 crash, after the 2000 crash, and the 2008 crash, the BSADF statistic actually decreases while the regime switching model sees explosive bubble periods.

5 Conclusion

In this paper, we propose a new regime switching model with two regimes, a normal regime and a bubble regime. To estimate the model we use a sequential Bayesian simulation method that allows for real time detection of bubble origination and conclusion. A particular feature of our framework is that it sequentially tracks the joint posterior distribution of the fixed parameters and of the hidden states. Hence, it properly allows for real time parameter uncertainty. The Monte Carlo evidence suggests that our method is reliable and quite robust to the presence of outliers and compares favorably to existing online methods in detection power. We carry out empirical study using real monthly S&P 500 price-dividend data. While some similar results have been obtained by PSY (2013a) in a classical setup and by our method, we find some differences in the two set of empirical results. In particular, our method detects more bubble periods and can better discriminate between bubbles and collapses.

References

- [1] Ahamed, L., 2009, *Lords of Finance: The Bankers Who Broke the World*, Penguin Press, New York.
- [2] Andrieu C, Doucet A and Holenstein A, 2010, Particle Markov Chain Monte Carlo, *Journal of Royal Statistical Society B*, 72,1-33.
- [3] Caballero, R.J., Farhi, E. and Gourinchas, P.O., 2008, Financial Crash, Commodity Prices and Global Imbalances, *Brookings Papers on Economic Activity*, 2, 1-55.
- [4] Godfeld S., Quandt, R., 1973, The Estimation of Structural Shifts By Switching Regressions, *Annals of Economic and Social Measurement*, 2, 473–483.
- [5] Cochrane, J, 2011, Presidential Address: Discount Rates, *Journal of Finance* 66, 1047–1108.
- [6] Chopin N, Jacob PE and Papaspiliopoulos O, 2013, SMC²: An efficient algorithm for sequential analysis of state space models, *Journal of the Royal Statistical Society B*, 75, 397-426.
- [7] Fearnhead, P, 1998, *Sequential Monte Carlo methods in Filter Theory*. PhD Thesis, Department of Statistics, University of Oxford
- [8] Fulop A and Li J, 2013, Efficient learning via simulation: A marginalized resample-move approach, *Journal of Econometrics*, 176, 146-161.
- [9] Homm U., Breitung, J., 2012, Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods, *Journal of Financial Econometrics*, 10, 198-231.
- [10] Funke, M., Hall, S.G. and Sola, M., 1994, Rational bubbles during Poland’s hyperinflation: implications and empirical evidence, *European Economic Review*, 38, 1257-1276.

- [11] Gouriéroux, C., Zakoian, J.M., 2013, Explosive Bubble Modelling by Noncausal Process, Working Paper, CREST
- [12] Hall, S.G., Psaradakis, Z. and Sola, M., 1999, Detecting Periodically Collapsing Bubbles: A Markov-Switching Unit Root Test, *Journal of Applied Econometrics*, 14, 143–154.
- [13] Hamilton, J.D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.
- [14] Kass, R.E., Raftery, A.E., 1996, Bayes factors, *Journal of the American Statistical Association*, 90, 773-795.
- [15] Phillips, P.C.B., Shi, S.P and Yu, J., 2013a, Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P500, Working Paper 1914, Cowles Foundation, Yale University.
- [16] Phillips, P.C.B., Shi, S.P and Yu, J., 2013b, Testing for Multiple Bubbles: Limit Theory of Real Time Detector, *International Economic Review*, forthcoming.
- [17] Phillips, P.C.B., Shi, S.P and Yu, J., 2014, Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behavior, *Oxford Bulletin of Economics and Statistics*, 76, 315-333.
- [18] Phillips, P.C.B., Wu, Y., and Yu, J., 2011, Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review* 52, 201-226.
- [19] Shi, S.P., 2013, Specification Sensitivities in the Markov-Switching Unit Root Test for Bubbles, *Empirical Economics*, 45, 697-713.
- [20] Shi S.P., Song Y. 2014, Identifying Speculative Bubbles with an Infinite Hidden Markov Model, Working Paper.

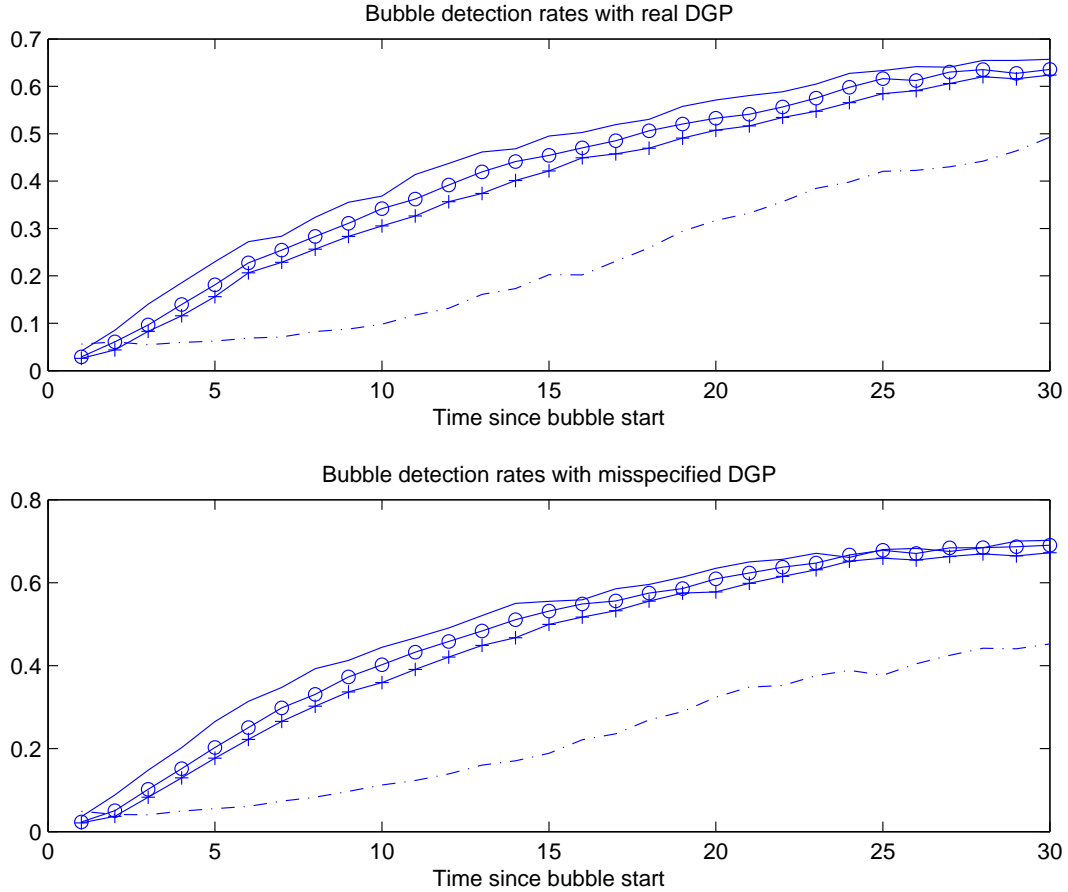
- [21] Yiu, M., Yu, J. and Jin, L. 2013, Detecting Bubbles in Hong Kong Residential Property Market, *Journal of Asian Economics*, 28, 115-124

Table 1: Monte Carlo Results on Posterior Parameter Estimates

	True	Post. Mean	Post. 90% Credible Interval	Prior 90% Credible Interval
λ_1	150	189.4	126.9	153.7
k_2	1.8	1.91	1.68	1.87
μ_2	30	35.95	21.15	31.01
z_{11}	0.98	0.97	0.01	0.89
z_{22}	0.94	0.93	0.049	0.90
σ_l	0.65	0.65	0.04	2.7
σ_m	2.8	2.77	0.37	1.88
δ	0.3	0.214	0.428	0.477
β_1	0.99	0.989	0.006	0.025
β_2	1.015	1.014	0.005	0.0180

This table reports various statistics from the Monte Carlo exercise using the true regime switching model as the data generating process, where we simulate 100 data sets with 1,698 observations. The first column reports the true parameter values, the second column the average posterior means across the 100 data sets, the third column the average length of the posterior central 90% credible interval across the 100 runs and the fourth column the prior 90% credible interval.

Figure 1: Monte Carlo Results on Bubble Detection Rates of Online Learning vs PSY (2013a)



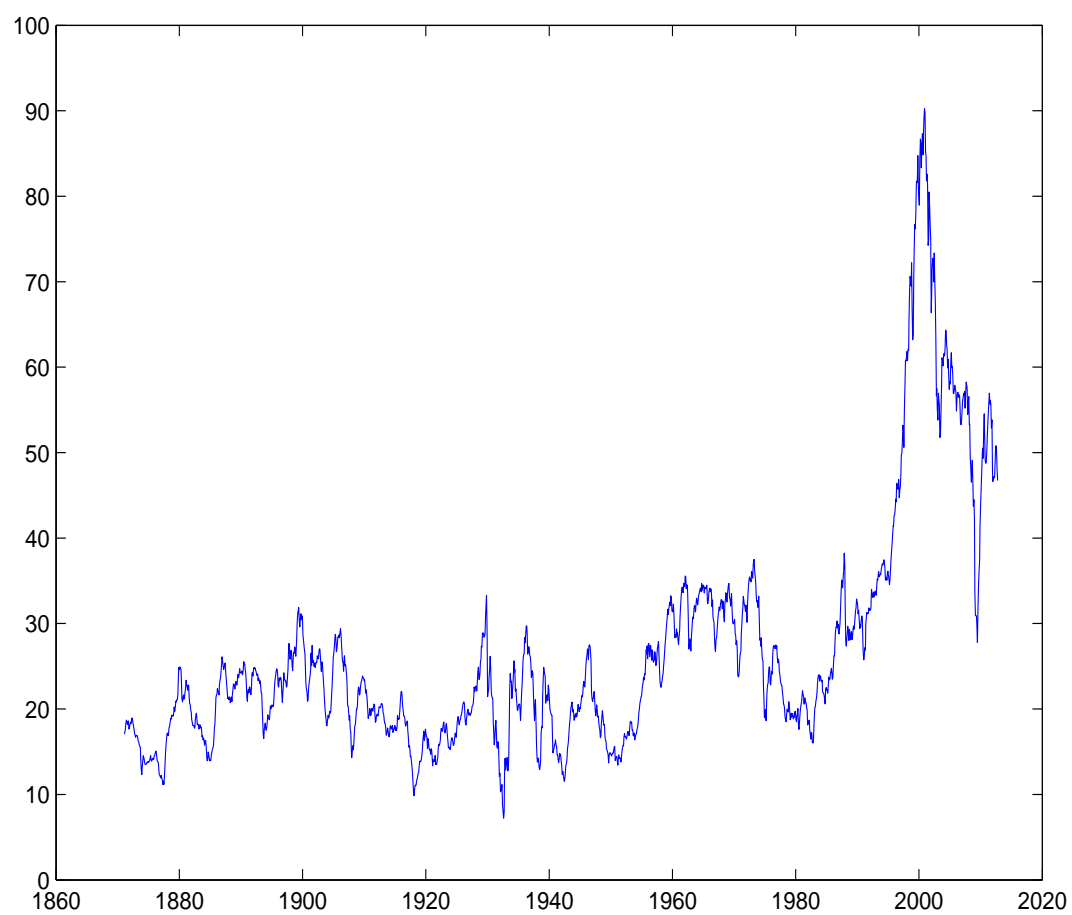
This table reports how well the online Bayesian filter and the PSY (2013a) method perform in detecting bubble regimes as a function of the duration of the bubble. We execute a Monte Carlo exercise using two data generating processes. First, we run simulations using the true regime switching model as the data generating process. Second, to investigate the effect of outliers, we simulate data from a misspecified version with student-t innovations with 4 degrees of freedom. For each DGP, we simulate 100 data sets with 1,698 observations. The upper panel shows the average frequency of periods flagged as bubble among the time periods that are in fact in the bubble regimes since n periods when the DGP from our regime switching model is used to generate the data. The solid lines depict the detection rates from our online Bayesian filter. The flat line corresponds to a stamping rule of $\zeta = 1$, the one with circles to $\zeta = 2$ while the one with crosses to $\zeta = 3$. The dotted line presents the detection rates from the PSY (2013a) method. The lower panel reports analogous results when a misspecified DGP with student-t innovations with 4 degrees of freedom is used.

Table 2: Monte Carlo Bubble Detection Statistics

	Number of Bubble Spells	Total Bubble Length/T	Avg Bubble Duration in months
<i>Panel A: Correctly Specified DGP</i>			
PSY	13.2	0.095	11.0
RS, $\zeta = 1$	17.4	0.095	8.9
RS, $\zeta = 2$	8.6	0.091	17.3
RS, $\zeta = 3$	6.9	0.087	20.6
<i>Panel B: Misspecified DGP</i>			
PSY	13.8	0.096	11.0
RS, $\zeta = 1$	19.9	0.106	8.7
RS, $\zeta = 2$	10.0	0.102	17.0
RS, $\zeta = 3$	8.0	0.099	21.0

This table reports summary statistics on the different bubble-stamping procedures in a Monte Carlo exercise using two data generating processes. First, we run simulations using the true regime switching model as the data generating process. Second, to investigate the effect of outliers, we simulate data from a misspecified version with student-t innovations with 4 degrees of freedom. For each DGP, we simulate 100 data sets with 1,698 observations. In both panels, the first row reports the results from the PSY (2013a) procedure while rows 2-4 report the results from our regime switching model at different values of the bubble stamping parameter ζ . In all cases the figures are average numbers across the 100 simulations.

Figure 2: S&P 500 Price-Dividend Ratio



This figure shows the monthly real S&P 500 P/D data between January 1871 to June 2012.

Table 3: Full Sample Posterior Parameter Estimates for S&P 500

	Posterior Mean	Posterior 5 th Prctile	Posterior 95 th Prctile
λ_1	147	123.5	183.1
k_2	1.795	1.152	2.589
μ_2	30.74	25.25	38.31
z_{11}	0.9842	0.9754	0.9907
z_{22}	0.9412	0.9128	0.963
σ_l	0.6694	0.6367	0.7017
σ_m	2.895	2.708	3.1
δ	0.3117	0.09392	0.6485
β_1	0.99	0.9784	0.9982
β_2	1.015	1.01	1.018

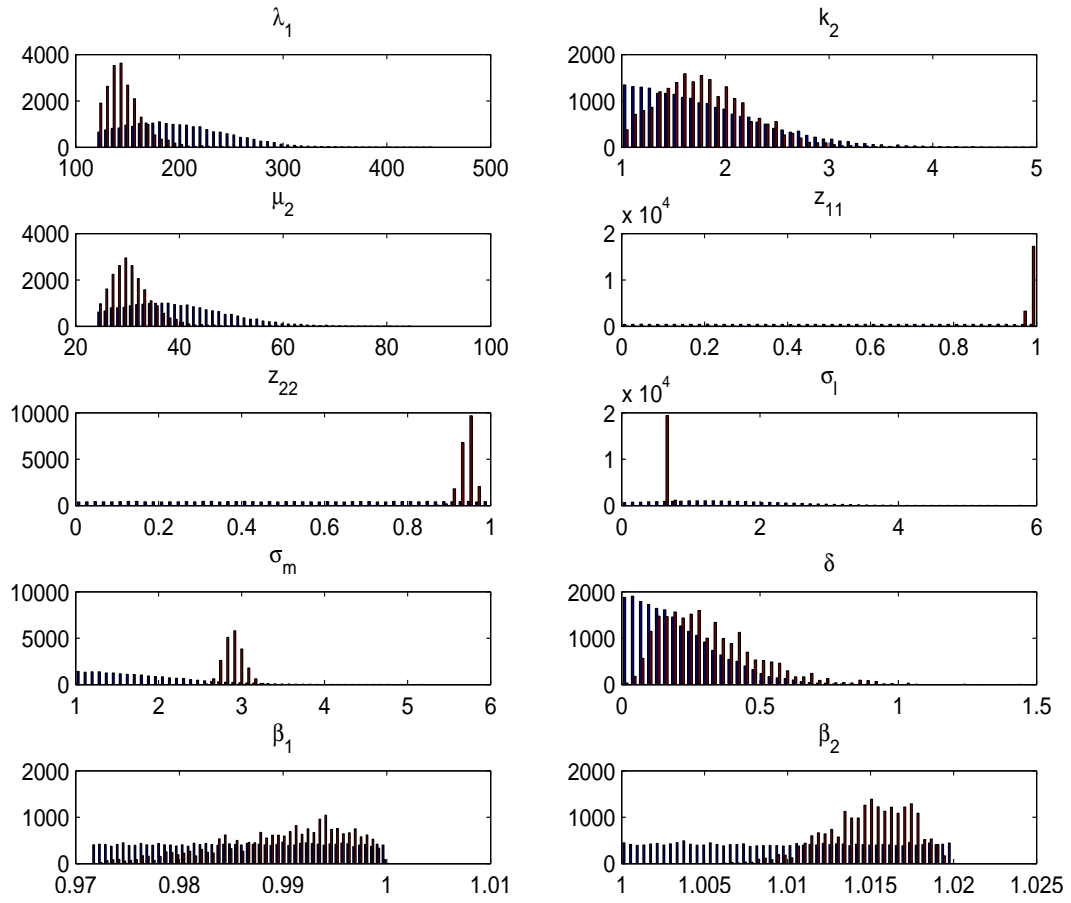
This table reports the full sample posterior estimates of the full model on monthly S&P 500 P/D data between January 1871 to June 2012.

Table 4: Bubble Detection Statistics for S&P 500

	Number of Bubble Spells	Total Bubble Length/T	Avg Bubble Duration in months
PSY	22	0.056	4.27
RS, $\zeta = 1$	58	0.16	4.5
RS, $\zeta = 2$	24	0.14	9.7
RS, $\zeta = 3$	20	0.125	10.4

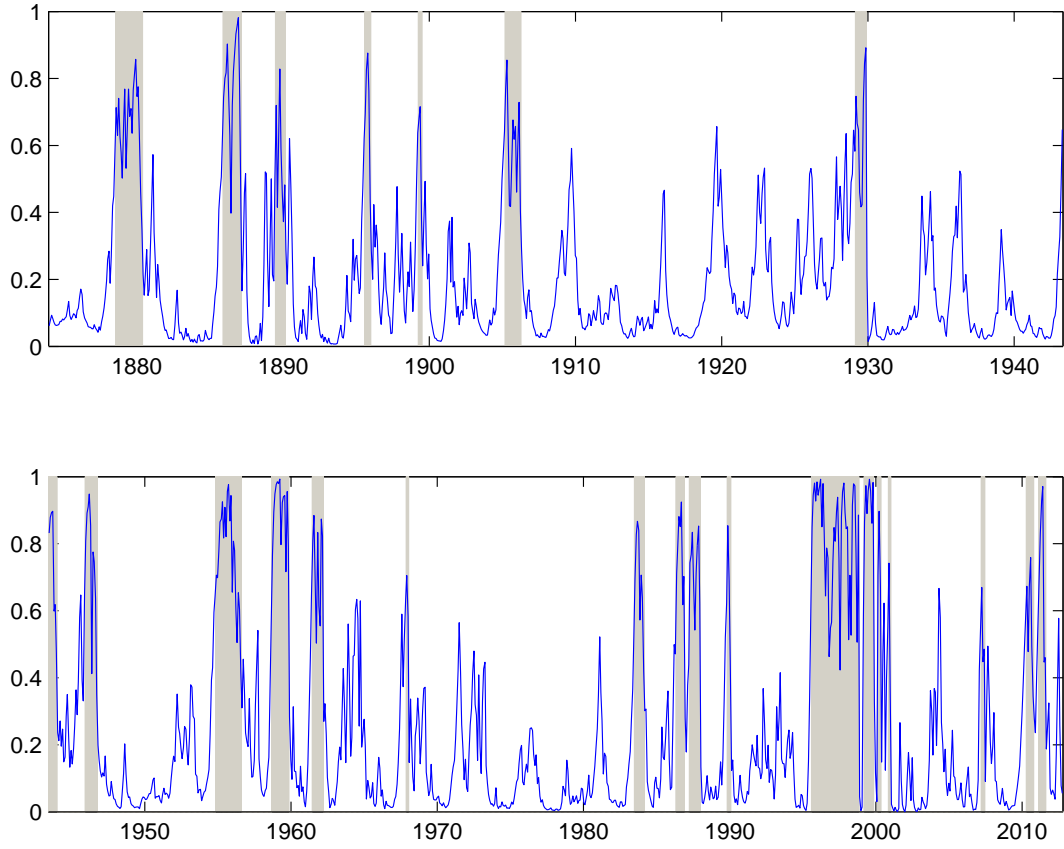
This table reports summary statistics on the different bubble-stamping procedures for monthly S&P 500 P/D data between January 1871 to June 2012. The first row reports the results from the PSY (2013a) procedure while rows 2-4 report the results from our regime switching model at different values of the bubble stamping parameter ζ .

Figure 3: Histogram of parameter priors and posteriors



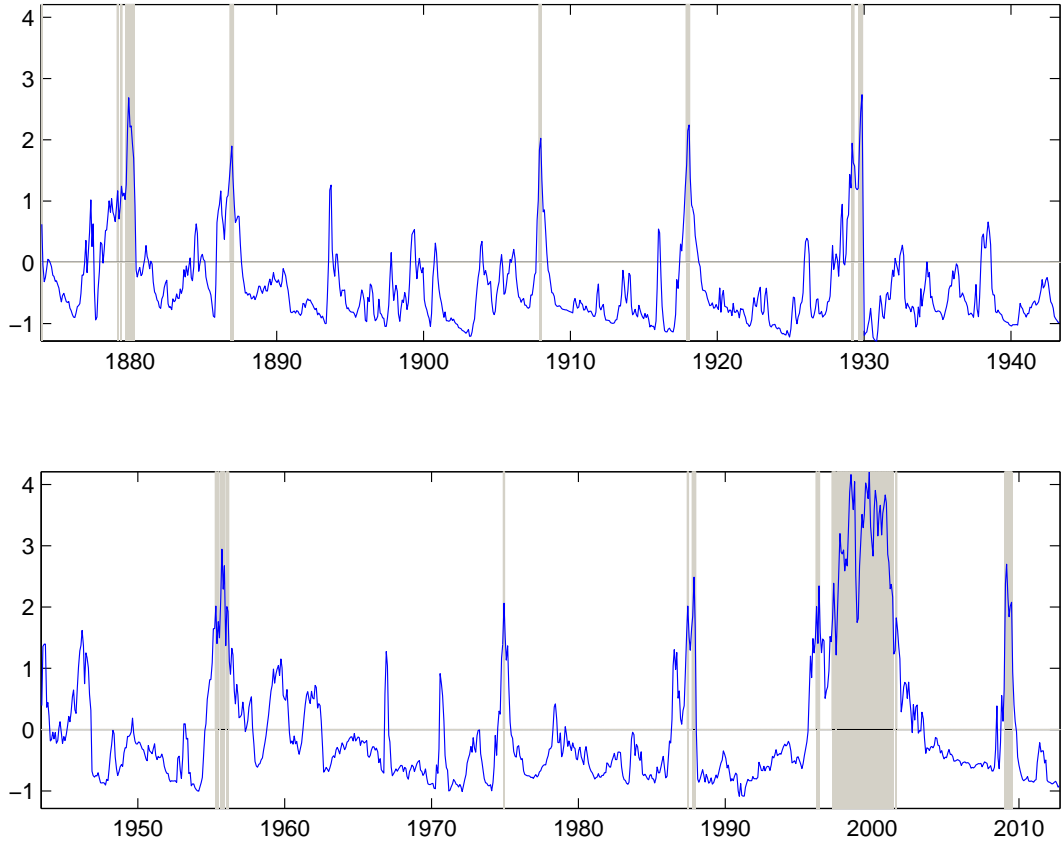
This figure reports the histogram of the priors (in blue) and the full-sample posteriors (in red). The sample is monthly S&P 500 P/D data between January 1871 to June 2012.

Figure 4: Bubble Regimes from Bayesian Learning



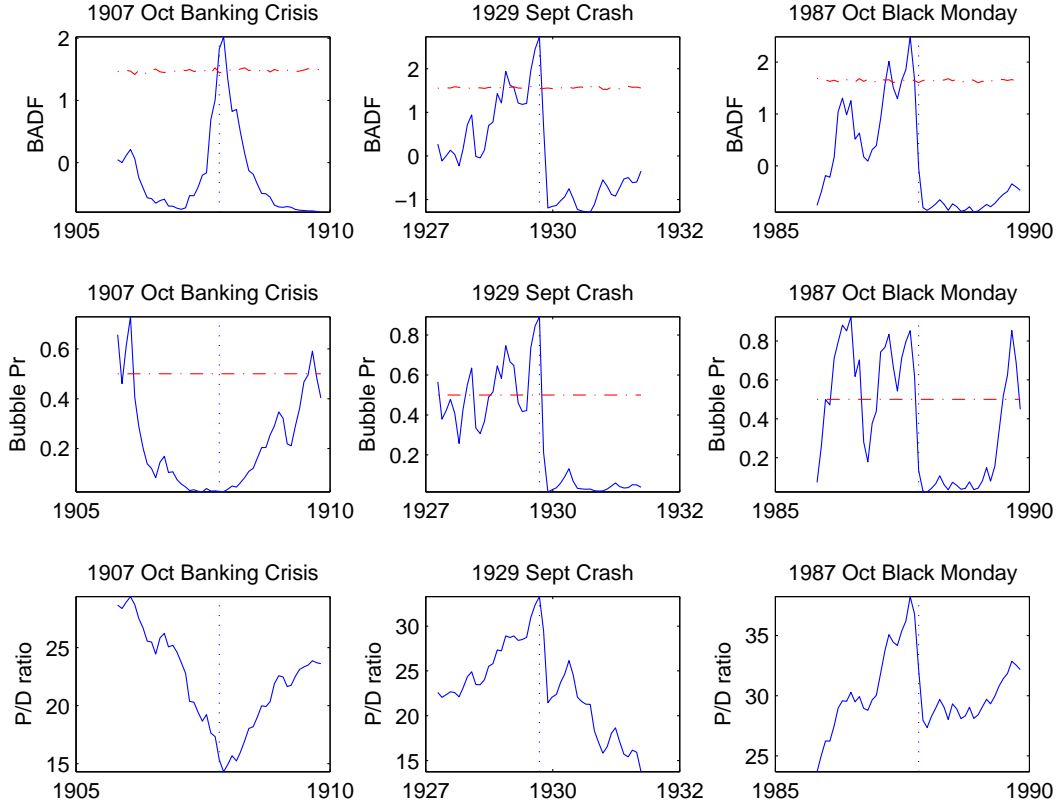
This figure reports the real-time bubble regimes indicated by our regime switching model together with the filtered bubble regime probability. A given month is classified as belonging to the bubble regime if this latter is the regime with the highest filtered probability. The plot corresponds to the bubble stamping parameter $\zeta = 2$. Bubble regimes are the shaded grey areas. The sample is monthly S&P 500 P/D data between January 1871 to June 2012.

Figure 5: Bubble Regimes from recursive regressions as in PSY (2013a)



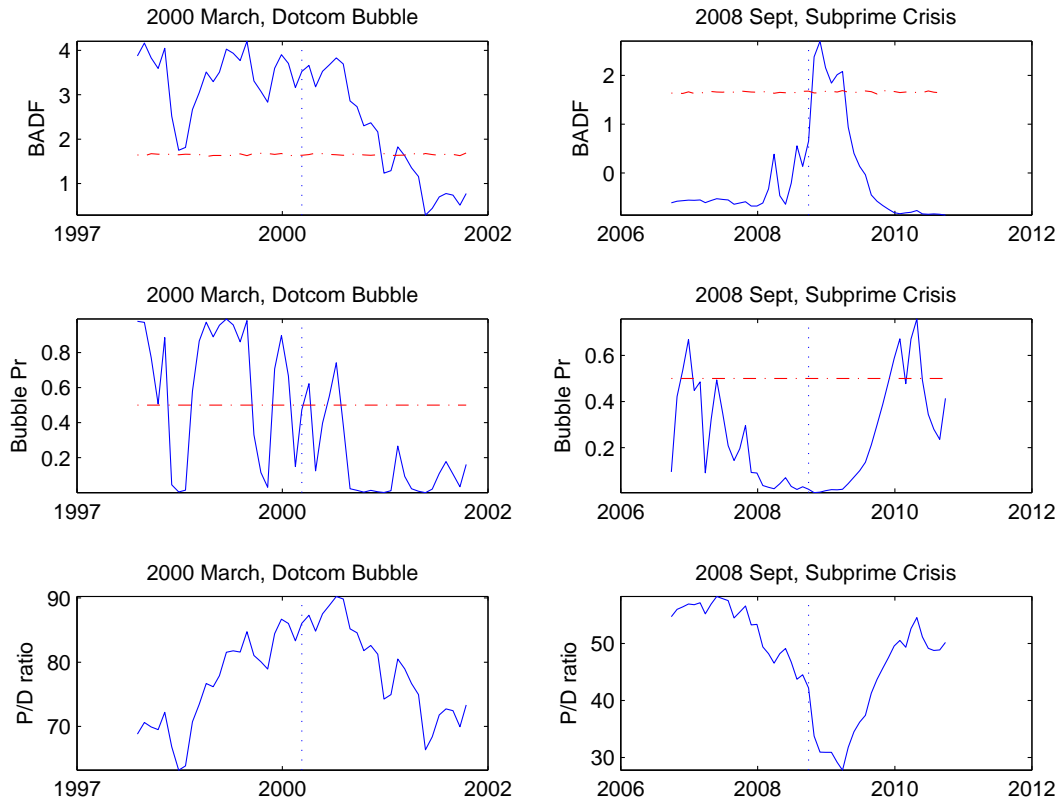
This figure reports the real-time bubble regimes indicated the backward sup ADF (BSADF) statistics from PSY (2013a). A given month is deemed to belong to a bubble regime if the value of the BSADF statistic exceeds $Q_{BSADF}(0.95) + \log(t)/100$ where $Q_{BSADF}(0.95)$ is the 95% critical value of the test statistic. Bubble regimes are the shaded grey areas while the solid line depicts the BSADF sequence. The sample is monthly S&P 500 P/D data between January 1871 to June 2012.

Figure 6: Behavior of bubble detectors around historical events 1



This figure reports the behavior of both the PSY and our regime-switching bubble indicators around some well-known historical events. The event itself is always shown by a vertical dashed line. Each time we report two years before and two years after the event. The first row reports the PSY BSADF statistic together with $Q_{BSADF}(0.95) + \log(t)/100$ in dashed red. The second row report the filtered bubble probabilities from our regime-switching model. The last row shows the data, the real S&P 500 P/D ratio.

Figure 7: Behavior of bubble detectors around historical events 2



This figure reports the behavior of both the PSY and our regime-switching bubble indicators around some well-known historical events. The event itself is always shown by a vertical dashed line. Each time we report two years before and two years after the event. The first row reports the PSY BSADF statistic together with the 95% critical values in dashed red. The second row report the filtered bubble probabilities from our regime-switching model. The last row shows the data, the real S&P 500 P/D ratio.